

The background of the slide features abstract, flowing green lines that create a sense of movement and depth. These lines are concentrated on the left side and curve towards the right, blending into a light green background.

# **Supervised Learning Algorithm**

**Regression : Random Forest**

By Gayatri Jondhale

# Contents

- What is Random Forest and History
- Definition
- Random Forest Model
- Why Use Random Forest Model
- Advantages and Disadvantages
- Random Forest Application

# Introduction

- A random Forest Algorithm is supervised machine learning Algorithm consisting decision trees.
- The general method of random decision forests was first proposed by Ho in 1995. After that, it was developed by Leo Breiman in 2001.

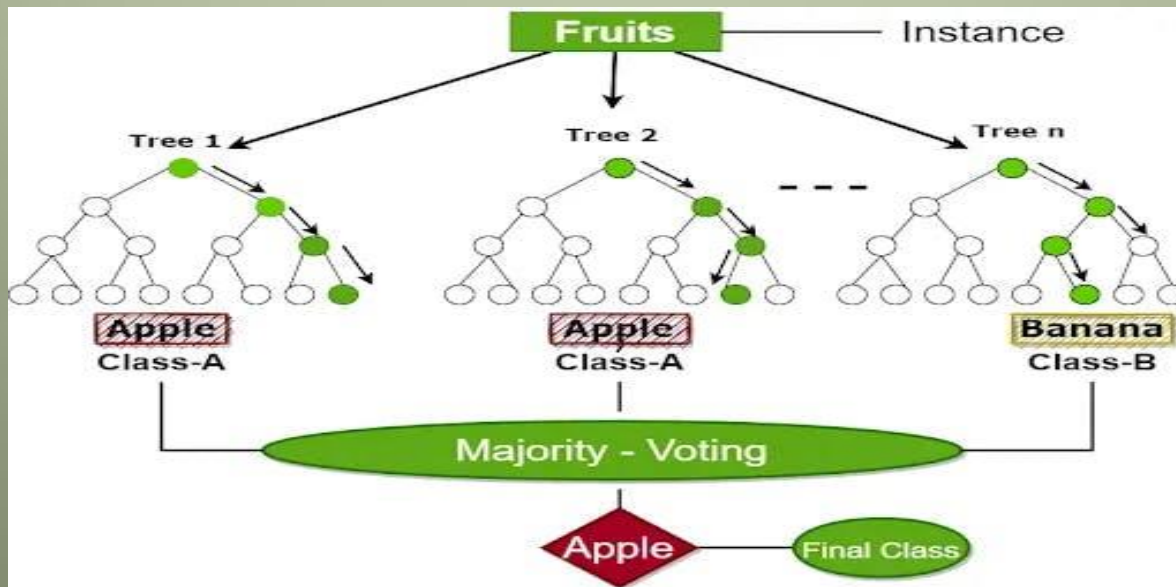
# Definition

Random Forest algorithm is a powerful tree learning technique in Machine Learning. It works by creating a number of Decision Trees during the training phase.

- Random forest or random decision forests is an ensemble learning method for classification, regression.
- Random decision forests correct decision tree habit of overfitting to their training set.

# Random Forest Model

- Create Bootstrap Dataset from Original data by randomly choosing data (repetition is allowed).
- Create Randomized Decision tree from Bootstrap dataset.
- Finally output of the random forests is the class selected by most trees.



# Why Use Random Forest Model

- Random forest are an effective tool in prediction.
- Forests give results competitive with boosting and adaptive bagging .yet do not progressively changes the training set.
- Random inputs and random features produce good results in classification- less in regression.
- For larger data sets. we can gain accuracy by combining random features with boosting.

# Advantages

- Versatile uses
- Easy-to-understand hyperparameters
- Classifier doesn't overfit with enough trees

# Disadvantages

- Increased accuracy requires more trees
- More trees slow down model
- Can't describe relationships within data



# Random Forest Applications

- Detects reliable debtors and potential fraudsters in finance
- Verifies medicine components and patient data in healthcare
- Gauges whether customers will like products in e-commerce